Comparison among Harpertown, Nehalem-EP and Westmere

Category: Pleiades

DRAFT

This article is being reviewed for completeness and technical accuracy.

Among the three processor types used in Pleiades, Nehalem-EP and Westmere are very similar to each other, while Harpertown is significantly different from the other two.

The main differences between Nehalem-EP and Westmere processors are:

- Both Nehalem-EP and Westmere have 24 GB of memory per node. However, there are 8 cores per Nehalem-EP node vs 12 cores per Westmere node, resulting in more memory per core for Nehalem-EP (3 GB/core) than Westmere (2 GB/core).

- The size of the L3 cache is 8 MB per quad-core for Nehalem-EP while it is 12 MB per 6-core for Westmere.

- For inter-node communication, two devices are involved: the Infiniband switches and the host channel adapter chip (HCA). For both the Nehalem-EP and Westmere racks, there are two SGI Infiniband QDR switches per half-IRU (which comprises 8 nodes). One of the switches is used for ib0 (used mainly for MPI communication), and the other for ib1 (used mainly for IO). The maximum raw data transfer rate through these switches is 40 Gigabits per second (Gbps). However, the HCA on each motherboard (one node per motherboard) is different for Nehalem-EP and Westmere. For Nehalem-EP, a 4x DDR HCA with a raw data transfer rate of 20 Gbps is used. For Westmere, a 4x QDR HCA with a rate of 40 Gbps is used. This difference results in better inter-node communication performance between the Westmere nodes than between the Nehalem-EP nodes.

  Note: The communication path between pairs of nodes vary, depending on where the nodes are relative to each other. For example:

  - two nodes on the same half-IRU: HCA to IB switch on the half-IRU to HCA.
  - two nodes on the same IRU but different half-IRUs: HCA to IB switch (of one half-IRU) to IB switch (of the other half-IRU) to HCA.

The main differences between Harpertown and Nehalem-EP/Westmere processors are:

- The processor labeling in Harpertown is not contiguous. On the contrary, the labeling in Nehalem-EP/Westmere is contiguous.

- Nehalem-EP/Westmere incorporates the SSE 4.2 SIMD instructions, which adds 7 new instructions to the SSE 4.1 set in Harpertown.

- Every two cores in Harpertown share a common L2 cache, while every core in Nehalem-EP/Westmere has its own private L2 cache. In addition, there is a L3 cache shared by the four cores in each socket of Nehalem-EP (or by the 6 cores in each socket of Westmere), while there is none for Harpertown.

- The Nehalem-EP based nodes have 3 GB/core (i.e., 24 GB/node) of memory as compared to 1 GB/core (i.e., 8 GB/node) in most of the Harpertown-based nodes in Pleiades.

  The Westmere based nodes have 2 GB/core (i.e., 24 GB/node) of memory as compared to 1 GB/core (i.e., 8 GB/node) in most of the Harpertown-based nodes in Pleiades.

- Nehalem-EP/Westmere, with a higher ratio of memory bandwidth to processor speed, is a better balanced system than the Harpertown.

  The key features which enable this improvement are the Intel QuickPath Interconnect, which provides communication with the other processor on the same node, and an integrated memory controller. Together they result in a higher aggregate bandwidth.

  In addition, each Nehalem-EP/Westmere core has its own L1 and L2 cache which helps to decrease the number of stalls in a data path. The data pre-fetch algorithm for L2 and L3 caches has been substantially reworked to achieve more effective data loads.

- Hyperthreading and TurboBoost are additional features on Nehalem-EP/Westmere, but not for Harpertown. Hyperthreading is available by user request for Nehalem-EP/Westmere. Turbo Boost is set to *ON* for Nehalem-EP/Westmere.

Article ID: 81
Last updated: 27 Oct, 2011
Computing at NAS -> Computing Hardware -> Pleiades -> Comparison among Harpertown, Nehalem-EP and Westmere
http://www.nas.nasa.gov/hecc/support/kb/entry/81/?ajax=1